

## **Model klasyfikacji, podział LOSOWY, ze stemmingiem**

Accuracy 0.7, F1 0.57, Precision 1, Recall 0.4, Specificity 1

### **Ocena wyników modelu**

Accuracy 70%: umiarkowana ogólna trafność modelu (30% wszystkich przypadków jest błędnie klasyfikowanych).

F1 Score 0.57: umiarkowany balans między Precision a Recall, nadal ograniczany przez niski Recall.

Precision 100%: model nie generuje fałszywych pozytywów (FP = 0). Każde przewidziane "tak" jest pewne.

Recall 40%: model wykrywa 40% rzeczywiście pozytywnych przypadków. Przewidziane "nie" nie są pewne.

Specificity 100%: Model poprawnie rozpoznaje wszystkie rzeczywiste „nie” (TN) i nigdy nie myli „nie” z „tak” (FP = 0).

### **Wniosek**

Umiarkowana ogólna trafność modelu.

Model jest umiarkowanie ostrożny i unika False Positives (Specificity = 100%), ale może jednocześnie popełniać False Negatives (nadal niski Recall) i pomijać większość rzeczywistych pozytywnych przypadków, co ogranicza jego użyteczność, mimo idealnej precyzji.

Model jest umiarkowanie dobry, ale wciąż wymaga poprawy w zakresie Recall, aby stać się praktycznym narzędziem wykrywania „tak”.

## **Model klasyfikacji, podział STRATYFIKOWANY, ze stemmingiem**

Accuracy 0.9, F1 0.75, Precision 1, Recall 0.6, Specificity 1

### **Ocena wyników modelu**

Accuracy 90%: wysoka ogólna trafność modelu (tylko 10% wszystkich przypadków jest błędnie klasyfikowanych).

F1 Score 0.75: dobry balans między Precision a Recall, model skutecznie równoważy pewność i wykrywalność.

Precision 100%: model nie generuje fałszywych pozytywów (FP = 0). Każde przewidziane "tak" jest pewne.

Recall 60%: model wykrywa 60% rzeczywiście pozytywnych przypadków. Przewidziane "nie" nadal mogą zawierać pominięte „tak” (40% pozostaje niewykryte).

Specificity 100%: Model poprawnie rozpoznaje wszystkie rzeczywiste „nie” (TN) i nigdy nie myli „nie” z „tak” (FP = 0).

### **Wniosek**

Wysoka ogólna trafność modelu.

Model osiąga wysoką trafność klasyfikacji (Accuracy = 90%). Model unika False Positives (Specificity = 100%) i nie popełnia błędnych klasyfikacji „tak” dla przypadków negatywnych. Dzięki poprawionemu Recall (60%), model unika False Negatives i wykrywa większość rzeczywistych pozytywnych przypadków, co czyni go praktycznym i zrównoważonym narzędziem klasyfikacyjnym.

Model jest godny zastosowania w praktyce, w automatycznym wykrywaniu treści recenzji klientów, którzy polecają skorzystanie z usług PLL LOT. Dzięki wysokiej Precyzji i Specyficzności model nie generuje fałszywych rekomendacji, co pozwala w pełni ufać recenzjom sklasyfikowanym jako Recommended = yes.

Jednocześnie, dzięki dobrej wykrywalności (Recall 60%), model wykrywa większość pozytywnych opinii, co czyni go praktycznym narzędziem do identyfikacji np. ambassadorów marki lub klientów o wysokim potencjale promocyjnym, będących autorami analizowanych recenzji.

## **Podsumowanie**

Model STRATYFIKOWANY ze stemmingiem jest lepszy niż LOSOWY ze stemmingiem, ponieważ przy zachowaniu pełnej Precyzji i Specyficzności, zapewnia wysoką ogólną trafność klasyfikacji tj. Accuracy 90%.

Model STRATYFIKOWANY ze stemmingiem wykrywa 3 na 5 rzeczywiście pozytywnych przypadków (Recall 60%), podczas gdy model LOSOWY wykrywa tylko 2 na 5 (Recall 40%).

Model STRATYFIKOWANY ze stemmingiem jest bardziej stabilny i bardziej godny zaufania, a także jest bardziej użyteczny w praktyce, ponieważ wykrywa znacząco większą liczbę istotnych przypadków.